WHAT IS CLAIMED IS:

1. A speech signal processing apparatus for performing speech synthesis by concatenating a plurality of selected synthesis units and modifying the synthesis units based on predetermined prosody parameters, said apparatus comprising:

distortion obtaining means for obtaining a distortion which may be generated from selection to synthesis of the synthesis units;

selection means for selecting synthesis units to be used for speech synthesis, based on the distortion obtained by said distortion obtaining means; and

speech synthesis means for performing speech synthesis based on the synthesis units selected by said selection means.

2. An apparatus according to Claim 1, wherein said selection means selects a plurality of synthesis units based on a phoneme series including a plurality of phonemes.

3. An apparatus according to Claim 1, wherein said distortion obtaining means obtains a distortion which may be generated in each of a plurality of synthesis units corresponding to one phoneme, and wherein said selection means selects one synthesis unit from among the plurality of synthesis units corresponding to the one phoneme.

4. An apparatus according to Claim 1, wherein said selection means selects the synthesis units to be used in speech synthesis so as to minimize

the distortion.

5. An apparatus according to Claim 1, wherein said distortion obtaining means obtains the distortion based on a concatenation distortion generated by concatenating a synthesis unit to another synthesis unit and a modification distortion generated by modifying the synthesis unit.

6. An apparatus according to Claim 1, wherein said distortion obtaining means uses a value obtained by adding a concatenation distortion generated by concatenating a synthesis unit to another synthesis unit and a modification distortion generated by modifying the synthesis unit as the distortion.

7. An apparatus according to Claim 3, wherein said distortion obtaining means calculates the distortion as a weighted sum of the concatenation distortion and the modification distortion.

8. An apparatus according to Claim 5, wherein said distortion obtaining means calculates the concatenation distortion using a cepstrum distance.

9. An apparatus according to Claim 5, wherein said distortion obtaining means calculates the modification distortion using a cepstrum distance.

10. An apparatus according to Claim 5, wherein said distortion

obtaining means includes a table storing modification distortions, and determines the modification distortion by referring to the table.

11. An apparatus according to Claim 5, wherein said distortion obtaining means includes a table storing concatenation distortions, and determines the concatenation distortion by referring to the table.

12. An apparatus according to Claim 1, further comprising:

input means for inputting text data;

language analysis means for performing language analysis of the text data; and

prosody-parameter generation means for generating the predetermined prosody parameters based on a result of analysis of said language analysis means.

13. A speech signal processing method comprising:

a distortion obtaining step of obtaining a distortion generated by concatenating a plurality of selected synthesis units and modifying the synthesis units based on predetermined prosody parameters;

a selection step of selecting synthesis units to be used for speech synthesis, based on the distortion obtained in said distortion obtaining step; and

a speech synthesis step of performing speech synthesis based on the synthesis units selected in said selection step.

14. A method according to Claim 13, wherein in said selection step, a

plurality of synthesis units are selected based on a phoneme series including a plurality of phonemes.

15. A method according to Claim 13, wherein in said distortion obtaining step, a distortion which may be generated in each of a plurality of synthesis units corresponding to one phoneme is obtained, and wherein in said selection step, one synthesis unit is selected from among the plurality of synthesis units corresponding to the one phoneme.

16. A method according to Claim 13, wherein in said selection step, the synthesis units to be used in speech synthesis are selected so as to minimize the distortion.

17. A method according to Claim 13, wherein said distortion obtaining means obtains the distortion based on a concatenation distortion generated by concatenating a synthesis unit to another synthesis unit and a modification distortion generated by modifying the synthesis unit.

18. A method according to Claim 13, wherein in said distortion obtaining step, a value obtained by adding a concatenation distortion generated by concatenating a synthesis unit to another synthesis unit and a modification distortion generated by modifying the synthesis unit is used as the distortion.

19. A method according to Claim 17, wherein in said distortion obtaining step, the distortion is calculated as a weighted sum of the

concatenation distortion and the modification distortion.

20. A method according to Claim 17, wherein in said distortion obtaining step, the concatenation distortion is calculated using a cepstrum distance.

21. A method according to Claim 17, wherein in said distortion obtaining step, the modification distortion is calculated using a cepstrum distance.

22. A method according to Claim 17, wherein in said distortion obtaining step, a table storing modification distortions is provided, and the modification distortion is determined by referring to the table.

23. A method according to Claim 17, wherein in said distortion obtaining step, a table storing concatenation distortions is provided, and the concatenation distortion is determined by referring to the table.

24. A method according to Claim 13, further comprising:

an input step of inputting text data;

a language analysis step of performing language analysis of the text data; and

a prosody-parameter generation step of generating the predetermined prosody parameters based on a result of analysis in said language analysis step.

25. A storage medium, capable of being read by a computer, storing a program for executing a method according to any one of Claims 13 through 24.